

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
7 juillet 2005 (07.07.2005)

PCT

(10) Numéro de publication internationale
WO 2005/062292 A2

(51) Classification internationale des brevets⁷ : G10L 13/08

(21) Numéro de la demande internationale :

PCT/FR2004/003278

(22) Date de dépôt international :

17 décembre 2004 (17.12.2004)

(25) Langue de dépôt :

français

(26) Langue de publication :

français

(30) Données relatives à la priorité :

0314928

18 décembre 2003 (18.12.2003)

FR

(71) Déposant (pour tous les États désignés sauf US) :

FRANCE TELECOM [FR/FR]; 6, place d'Alleray,
F-75015 Paris (FR).

(72) Inventeur; et

(75) Inventeur/Déposant (pour US seulement) : LASSALLE,
Edmond [FR/FR]; 17, rue du Hingar, F-22300 Lannion
(FR).

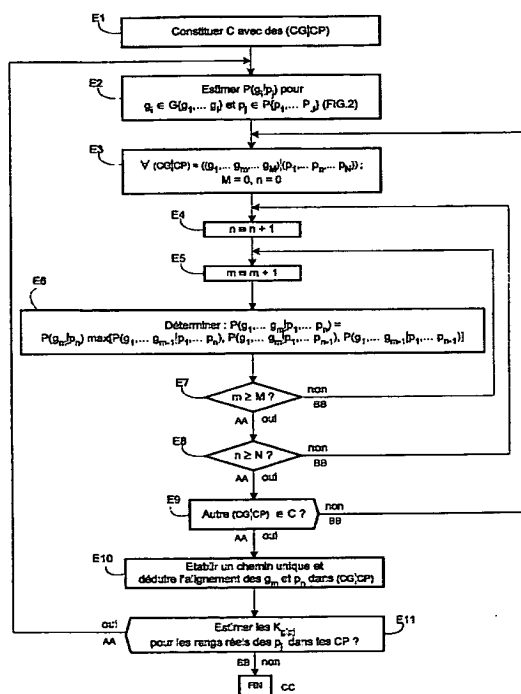
(74) Mandataire : CABINET MARTINET & LAPOUX;
43, boulevard Vauban, BP 405 Guyancourt, F-78055 Saint
Quentin Yvelines Cedex (FR).

(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,
MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,

[Suite sur la page suivante]

(54) Title: METHOD FOR AUTOMATIC CORRESPONDENCE BETWEEN GRAPHICAL AND PHONETIC ELEMENTS

(54) Titre : PROCEDE DE CORRESPONDANCE AUTOMATIQUE ENTRE DES ELEMENTS GRAPHIQUES ET DES ELE-
MENTS PHONETIQUES



E1 CONSTITUTE C WITH (CG)CP
E2 ASSESS $P(g_i, p_j)$
E3 DETERMINE
E4 OTHER
E5 ESTABLISH A SINGLE PATH AND DERIVE THE G_m AND P_n ALIGNMENT IN (CG)CP
E6 ASSESS K_{CP} POUR REAL RANGS OF P_n IN CP
AA YES
BB NO
CC END

(57) Abstract: The invention relates to a method for automatically segmenting any graphical chain into graphemes and any phonetic chain into phonemes by rewriting graphical chains (words) into phonetic chains. The inventive method consists in assessing first probabilities of the rewrite of graphical elements into phonetic elements (E2), in determining (E3-E9) second probabilities ($P(g_1, g_m/p_1, p_n)$) of MN rewrites of M graphical chains successively concatenating M graphical elements into N phonetic chains successively concatenating N graphical elements for each rewrite of a given graphical chain having M graphical elements into a corresponding phonetic chain having N graphical elements and in establishing links between the last elements (g_m, p_n) of the graphical and phonetic chains of the second rewrites in order to constitute, a path segmenting the given graphical chain into graphemes corresponding to phonemes segmenting the corresponding phonetic chain, respectively in the matrix MN of the second probabilities.

(57) Abrégé : L'invention déduit de transcriptions de chaînes graphiques (mots) en des chaînes phonétiques, une segmentation automatique de toute chaîne graphique en graphèmes et toute chaîne phonétique en phonèmes. Des premières probabilités de transcriptions des éléments graphiques en les éléments phonétiques sont estimées (E2). Pour chaque transcription d'une chaîne graphique donnée à M éléments graphiques en une chaîne phonétique correspondante à N éléments phonétiques sont déterminées (E3 - E9) des deuxièmes probabilités ($P(g_1, ..., g_m, p_1, ..., p_n)$) de MN transcriptions de M chaînes graphiques concaténant successivement les M éléments graphiques en N chaînes phonétiques concaténant successivement les N éléments phonétiques. Des liens sont établis entre les derniers éléments (g_m, p_n) des chaînes graphiques et phonétiques de deuxièmes transcriptions pour constituer dans une matrice MN des deuxièmes probabilités un chemin segmentant la chaîne graphique donnée en graphèmes

correspondant respectivement à des phonèmes segmentant la chaîne phonétique correspondante.



PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclarations en vertu de la règle 4.17 :

- relative au droit du déposant de revendiquer la priorité de la demande antérieure (règle 4.17.iii)) pour la désignation suivante US
- relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement

Publiée :

- sans rapport de recherche internationale, sera republiée dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.